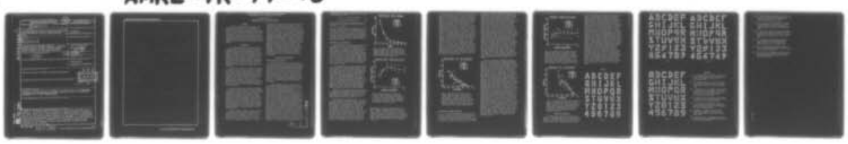
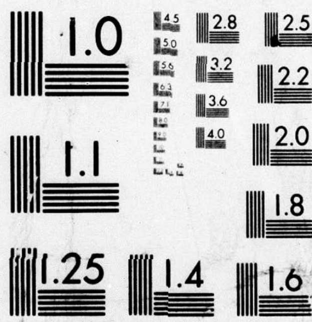
AD-A041 649 AEROSPACE MEDICAL RESEARCH LAB WRIGHT-PATTERSON AFB OHIO F/G 5/8
A SPATIAL FREQUENCY ANALYSIS MODEL FOR PREDICTING HUMAN PERFORM--ETC(U)
JUN 77 M W CANNON
UNCLASSIFIED AMRL-TR-77-43 NL

| OF |
AD
A041649



END
DATE
FILMED
8-77



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

12

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER (14) AMRL-TR-77-43	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) (6) A SPATIAL FREQUENCY ANALYSIS MODEL FOR PREDICTING HUMAN PERFORMANCE AT VISUAL PATTERN MATCHING TASKS		5. TYPE OF REPORT & PERIOD COVERED Proceedings
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) Mark W./ Cannon, Jr		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS AEROSPACE MEDICAL RESEARCH LABORATORY, AEROSPACE MEDICAL DIVISION, AIR FORCE SYSTEMS COMMAND, WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62202F, 7233-05-20 (16) (17)
11. CONTROLLING OFFICE NAME AND ADDRESS (12) 7p.		(11) 12. REPORT DATE June 1977
		13. NUMBER OF PAGES 7
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Prepared for presentation at the meeting and publication in the Proceedings of Systems, Man and Cybernetics Society, sponsored by IEEE, to be held in Washington, D. C., 19-21 September 1977.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A model for simulating human performance in visual pattern matching tasks is presented. The model is based on evidence of spatial frequency processing in the visual system, and on the hypothesis that shape recognition is determined only by the low spatial frequency harmonics of the image. Two psychophysical pattern matching experiments are described that demonstrate a clear functional relationship between the "similarity" of two patterns as judged by human observers and the Euclidean distance between spatially filtered Fourier transforms of the patterns.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

009 850

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

A SPATIAL FREQUENCY ANALYSIS MODEL FOR PREDICTING HUMAN PERFORMANCE AT VISUAL PATTERN MATCHING TASKS

Mark W. Cannon, Jr.
Aerospace Medical Research Laboratory
Wright-Patterson Air Force Base, Ohio

Summary

A model for simulating human performance in visual pattern matching tasks is presented. The model is based on evidence of spatial frequency processing in the visual system, and on the hypothesis that shape recognition is determined only by the low spatial frequency harmonics of the image. Two psychophysical pattern matching experiments are described that demonstrate a clear functional relationship between the "similarity" of two patterns as judged by human observers and the Euclidean distance between spatially filtered Fourier transforms of the patterns.

Introduction

The research described in this report represents a portion of the work being conducted at the Aerospace Medical Research Laboratory to develop quantitative models for observer-display interactions. These models will lead to design of displays optimally matched to human information processing capabilities under a variety of conditions. This paper addresses, in particular, the problem of predicting the confusability of symbols of the type that may be used in a graphic display. Alphanumeric symbols are used for the tests discussed here but the technique is not limited to these. Any two-dimensional display symbols can be analyzed by this technique.

Two bodies of research have helped to lead our work in its present direction. The first of these is the increasing amount of psychophysical and neurophysiological literature demonstrating the organization of the visual system as a spatial frequency analyzer. These works range from the early reports of Campbell et al.[1,2,3] to same of the more recent works by Hamerly, Quick and Reichert[4] and by Carlson, Cohen and Gorog[5]. These latter papers demonstrate clearly that from threshold to contrasts of at least 40%, the frequency analysis properties of the visual system can be closely approximated by linear mathematics.

The second body of work which has influenced our research is the application of spatial frequency analysis to the recognition of two-dimensional images by Kabrisky[6,7] and his students at the Air Force Institute of Technology. In a number of Masters theses[8] and Doctoral dissertations[9] over the past 10 years this group has demonstrated the machine recognition of printed characters can achieve a certain amount of font independence if only the low spatial frequencies (out to about the 3rd harmonic of the character width) are used in both prototypes and test characters. A simplified model of static shape recognition in the visual system that evolves from the synthesis of these two bodies of work is described below.

Images of objects (inputs) are formed on the retina and transmitted into the visual system via the optic nerve. At some point, (perhaps even in the retina) a two-dimensional spatial frequency transformation is performed on the input image. Many subsequent cognitive processes have access to all spatial frequency components, but the process that identifies what object is present requires only the low spatial harmonics of the image. In this identification process, the low-pass filtered input image is treated as a multidimensional vector. This vector is compared with a set of stored prototype vectors derived from a set of previously learned low-pass filtered images. The input vector is identified as belonging to that class of objects represented by the nearest prototype. (Nearest is defined as the shortest Euclidean distance in the pattern space coordinates.) Verification of this model requires the demonstration of a functional relationship between Euclidean distance derived from model predictions on a set of patterns with similarity judgements or misclassification probabilities derived from psychophysical experiments involving the same set of patterns. We have been able to demonstrate such a functional relationship using sets of alphanumeric characters as our test patterns. The remainder of this paper is devoted to describing two types of psychophysical tests and associated simulation results which provide a strong partial validation of the model.

Description of Experiments

Computer Analysis of Symbols

A description of the computer processing performed on the symbol sets is in order at this point, since essentially the same model distance predictions are used to analyze both psychophysical experiments. The symbols used were digitized as ones on a background of zeros for computer analysis. Each symbol, except for a few in set 1, has a maximum size of 10 x 14 points, and is located in a 32 x 32 background window of zeros. The four symbol sets used in these experiments are shown in the appendix in the same form in which they were presented to the human observers. The sawtooth effect of digitized diagonal lines was therefore presented to both humans and computer.

In both experiments, the same symbol set was used as both input and prototype. As we will see, this is fully justified in experiment 1, and is the best we can do at present for experiment 2. The 36 symbols were Fourier transformed using a two-dimensional Cooly-Tukey FFT algorithm and filtered using a square low-pass filter. The dc term and all harmonic components up to the maximum desired were saved while all higher terms were set equal to zero. Next, a 36 x 36 correlation matrix was generated. Each row of this matrix contained the maximum value of the cross-correlation functions computed between the filtered input symbol and each of the other symbols of the set. The correlation function for a symbol pair (i,j) was determined by multiplying the filtered spectrum of symbol i with the complex conjugate of the filtered spectrum of symbol j and taking the inverse transform of the product. The maximum amplitude of this inverse transform was the maximum correlation coefficient. Finally, a distance matrix D was computed from the correlation matrix by determining the Euclidean distances at which maximum correlation occurred. This distance is the minimum mean square distance between the two patterns, and can be derived from the maximum correlation coefficient, as we now show. Let X and Y be the position vectors of

A1

two patterns in the spatial domain and let d be the distance between X and Y.

$$d = |X - Y| = \sqrt{(X - Y) \cdot (X - Y)} \qquad (1)$$

$$d = \sqrt{(X \cdot X - 2X \cdot Y + Y \cdot Y)} \qquad (2)$$

The vectors have been energy-normalized after filtering so the dot products are equal to 1 and the distance is

$$d = \sqrt{2 - 2\rho} \qquad (3)$$

where ρ is the cross-correlation between X and Y. If we let ρ be the maximum of the cross-correlation function, the distance d is minimized. The value of d can range from zero to 1.4, since ρ ranges from zero to 1. The 36 x 36 distance matrix D contained these minimum distances between each input symbol and all other symbols of the set. Diagonal elements were zero, since these represent the distance from each symbol to itself.

Experiment 1: Shape Matching

The subjects were seated before a chart containing symbols of font 1, 2 or 3, as shown in the appendix, but the symbols on the experimental charts were arrayed in a random order. The relative distances between the symbols was larger than that shown on the charts in the appendix and the symbols themselves were 2.5 cm wide. The subject was seated at a distance from the chart such that the symbol width was one degree of visual angle. Subjects were given a randomized list naming all 36 symbols in the set. The subjects were instructed to locate on the chart a symbol named in the random list and then to report which other symbol on the chart matched it most closely in shape.

In the account below, we will refer to the symbol to be matched as the test symbol and the symbol chosen to match it as the comparison symbol. Set 1 was viewed by 35 subjects; set 2 by 23 subjects and set 3 by 27 subjects. The results showed that for each symbol of the test set, there was a good deal of agreement about which comparison symbol was closest. If the comparison symbols for each test symbol are arranged in rank order by the number of times it was chosen as closest, and the number in each rank is summed over all three sets as in Fig. 1, we see that approximately 50% of the choices are in rank 1. In fact, the curve is nearly perfectly exponential by rank. In rank 2, 25% of the people agree on the closest match and in rank 3 approximately 12% agree that this is the closest match.

Since 75% of the subjects agree that the rank 1 or 2 choices are the closest matches, we decided that our computer model would be given credit for a correct choice if the comparison symbol picked by the model as closest to each test symbol agreed with either the human rank 1 or 2 choice.

The results of the computer choices for a range of filter bandwidths is shown in Fig. 2. The ordinate gives the number of times the model agreed with human predictions. The best score across all three sets occurs at the eighth harmonic of the 32 x 32 viewing window, which is between the second and third harmonic of the symbol. The score at the sixth harmonic of the window, which is equal to the second harmonic of the symbol, is just marginally lower, so maximum performance occurs in the sixth to eighth harmonic range. The overall score is an 80% correct match with the human data. Note also that performance deteriorates at both larger and smaller filter bandwidths. We see that those symbol pairs with smallest intersymbol distances as

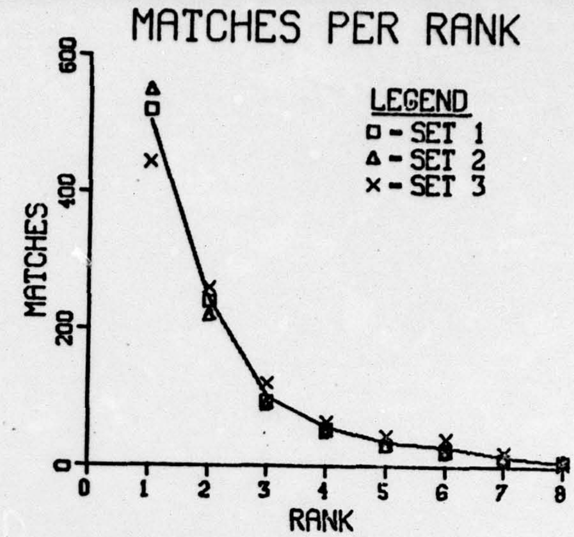


Fig. 1. A symbol rank is determined by how many subjects chose it as a closest match to the test symbol. This figure shows how many subjects chose symbols in each rank averaged across all symbols in the set. The solid line connects the average of each rank.

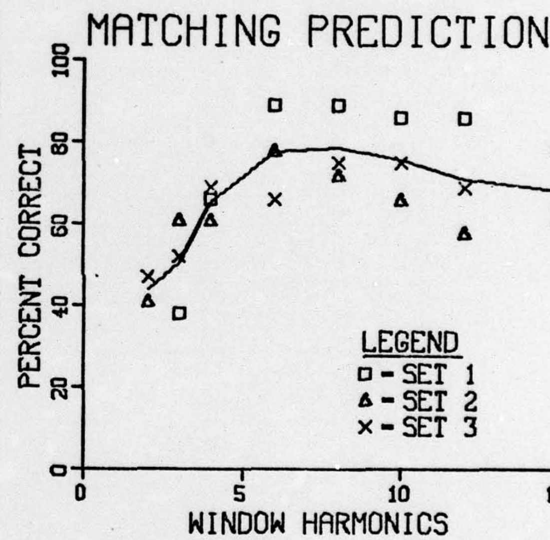


Fig. 2. The sample points represent percentage of model choices for closest match to a test symbol that agreed with either rank 1 or 2 choices of human sibjects. These are plotted as a function of the filter bandwidth as explained in the text. The solid line is an average of the sample points.

determined by our model correspond quite well with human judgments of symbol pairs most similar in shape. However, if shape similarity is related to intersymbol distance, there should be a functional relationship between the number of times a comparison symbol is chosen as most similar to a test symbol and the Euclidean distance between test and comparison symbols. This relationship is derived in the following way.

The human matching results for a particular symbol set are arranged into a 36 x 36 choice matrix C. Each entry C_{ij} is the number of times that subjects chose a comparison symbol j as most similar to test symbol i. The intersymbol distances corresponding to each i,j pair are contained in previously computed distance matrix D for this symbol set. A filter corresponding to the sixth harmonic of the window was used to compute D. Let us now divide up the distance axis into bins of width .05. Using matrices C and D, we add up all choices that fall in a given distance bin and divide by the number of test symbols that generated choices in that bin. This gives us the average number of choices per test symbol at a given distance between test and comparison symbols. Further normalization was accomplished for each set by dividing these averages by the number of subjects who took part for that set. The final averages for all three sets are plotted as a function of distance in Fig. 3. The points from the three sets overlap in a very *satisfactory manner*, almost as if all were derived from the same function. Thus Fig. 3 demonstrates, as we had hoped, that the relative number of times a particular comparison symbol is chosen to be a best match for a given test symbol decreases with increasing distance between comparison and test symbols. Apparently, similarity and Euclidean distance are functionally related for our spatially filtered symbols.

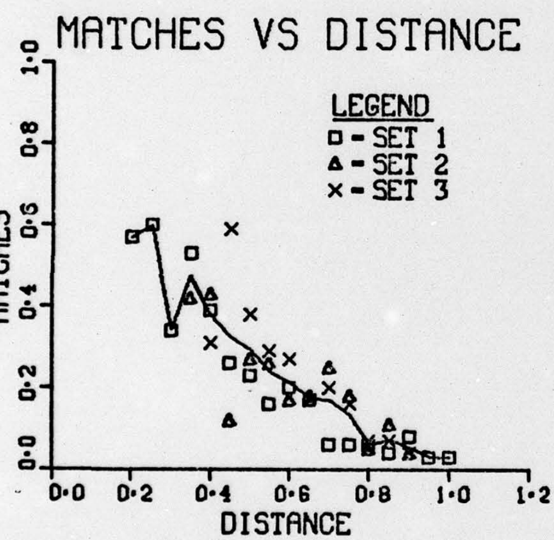


Fig. 3. The sample points represent the average number of times per test symbol that a given comparison symbol was picked as a closest match in shape to that test symbol. The abscissa is the model-predicted Euclidean distance between spatially filtered test and comparison symbols. The solid line is the average of the sample points, and shows a definite decrease with distance between symbols.

Experiment 2: Symbol Recognition

In a series of psychophysical experiments conducted at the Air Force Flight Dynamics Laboratory, Dr. Larry Goble has evaluated the confusability of three of the symbol sets shown in the appendix. Sets 2, 3 and 4 were used in his experiment. The subjects viewed the sets statically to become familiar with the shapes. They were then asked to identify the symbols when they were flashed on a screen and partially masked by a preceding and following uniform field of the same intensity as the symbol. The paradigm proceeded as follows: uniform field 10 msec, blank screen 5 msec, symbol 10 msec, blank screen 5 msec and uniform field 10 msec. This paradigm involves some significant differences from the matching experiment covered in the previous section. First, the short duration of the symbol presentation and the masking involve some temporal input parameters not yet treated in the model. We cannot yet say what effects stimulus duration would have on the complicated spectrum of a symbol, and we cannot adequately define the effect of pre- and post-stimulus masking. We will just assume these effects are small when we apply the model to analyze the results of these psychophysical experiments. The second problem is the prototype to which the test symbols are compared. We assume that the subject can learn the symbol shapes for each set well enough to compare the inputs to them. However, the prototype may be some combination of a wide variety of fonts to which the subject has been exposed. The fact that the subject can change his prototype set was demonstrated by Goble's data. Each subject viewed each symbol set 48 times and set 4 showed a distinct learning curve measured in number of correct responses. The other two sets showed trial-to-trial variation, but the average number of correct responses remained relatively constant over all trials.

The first comparison of model performance to human data was very similar to the comparison made in Fig. 2. The symbol with the smallest intersymbol distance in the ith row of the distance matrix D was picked as the model's choice for the symbol most confusable with the ith test symbol. We asked, "In how many cases does this choice agree with the symbol that produced either the largest or second largest number of human errors in each row?" The results are plotted in Fig. 4. The plot shows the number of *test symbols* for which model predictions agreed with human data as a function of the filter *bandwidth*. The striking difference between these plots and those of Fig. 2 is that they reach a maximum percentage of agreement at a filter bandwidth equal to the fifth harmonic of the window and are essentially flat above that bandwidth. Note also that the average curve has a maximum of only 47% compared to 80% for the static matching experiment. However, scanning the data showed us that there was still considerable correlation between the distribution of errors and Euclidean distance. In Fig. 5 we derive a curve giving the average number of human errors per test symbol as a function of the model-generated distance between the test symbol and a comparison symbol. We again divided the distance axis into bins of width .05. We summed the number of errors in each bin and divided by the number of test symbols for which errors were generated in that bin and by the total number of errors for all symbols of that set. The total errors varied considerably over the three sets. There were 1639 errors for set 2, 3184 errors for set 3 and 2893 errors for set 4. The averages for each set were plotted as the sample points in Fig. 5, and we see that there is a definite decrease in number of errors per test symbol as the distance between test and comparison symbol increases. Again, all three symbol sets give results which fall along the same curve, independent of the symbol set, depending only on Euclidean distance.

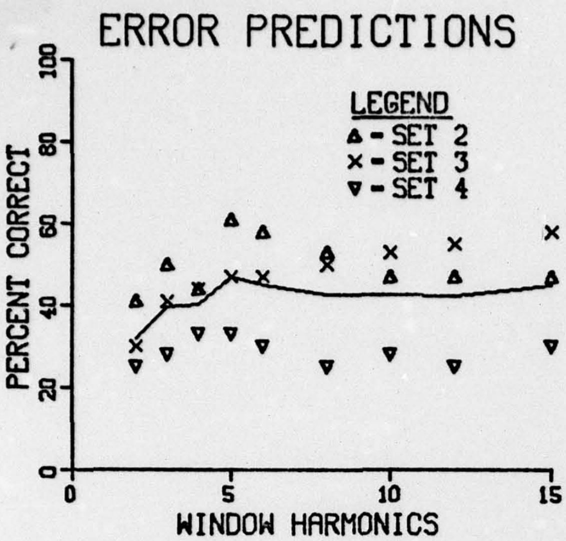


Fig. 4. The sample points represent the percentage of times that the symbol pairs with smallest model-derived intersymbol distances were the same as those symbol pairs having the largest or second largest number of confusions in a symbol identification experiment. The solid line is the average of the points. The abscissa gives the bandwidth of the model spatial frequency filter as explained in the text.

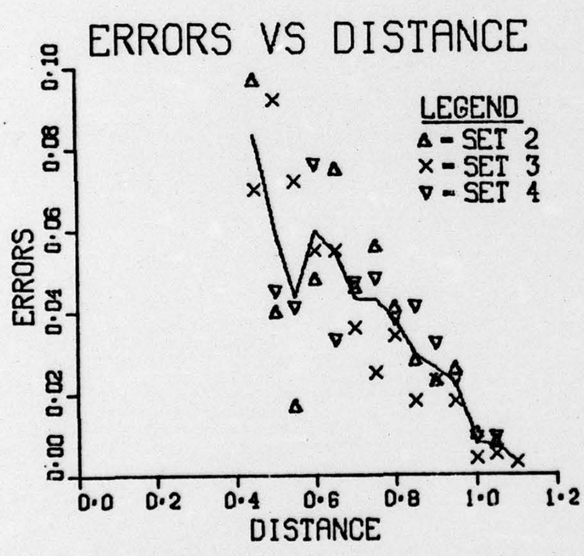
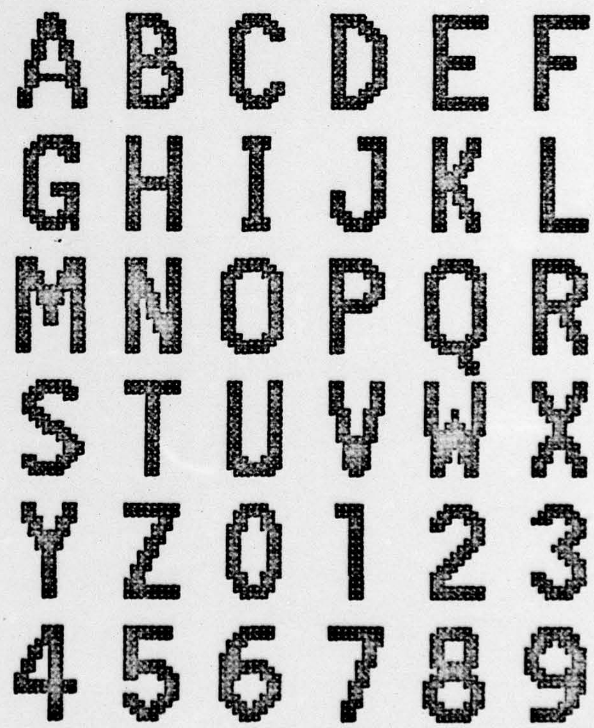


Fig. 5. The sample point ordinates represent the number of times per test symbol that a test symbol was mistaken for a particular comparison symbol. The solid line is the average of the sample points and demonstrates a decrease in the number of classification errors with an increase in the model-predicted distance between the symbols.

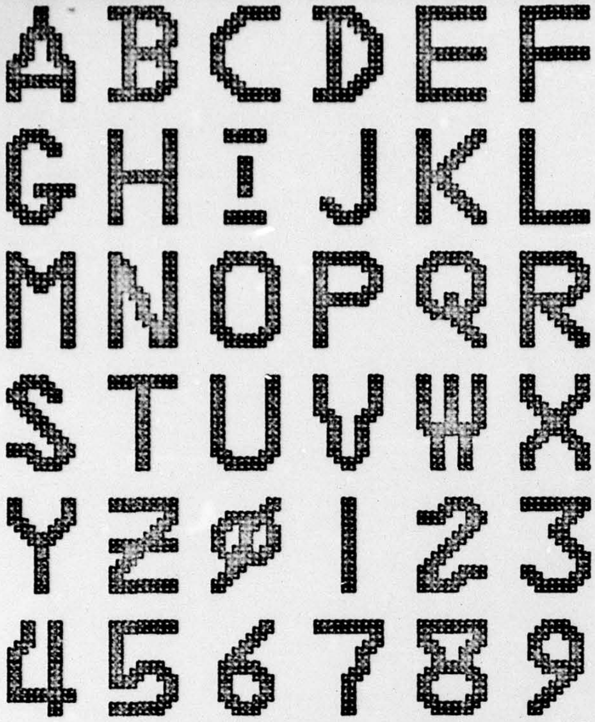
Discussion of Results

The model accounts very well for human performance at a pattern-matching task involving static imagery. Human shape-similarity judgments can be modeled by minimum Euclidean distances provided that the imagery is static. The model does not do so well in attempting to predict which particular symbol pairs will produce the most errors (Fig. 4) when the symbol images are dynamic. The general trends which were shown for the static experiment do hold, however, and we may say that on the average those symbol pairs which are shown to be close together by the static analysis model will produce more confusions in the dynamic case than those symbol pairs shown to be far apart statically. It may be possible to treat the uniform masking fields as some sort of noise in the system, since it is their presence that causes most of the errors. A number of researchers[10,11,12] have developed models of temporary visual memory storage which may be applicable. Sperling, in particular, has demonstrated that presentation of a uniform pre- and post-stimulus masking field interferes with the storage of symbols in this temporary memory area. However, a much more general model of spatio-temporal interactions in the visual system is required before we can apply this promising approach to the analysis of dynamic display imagery. The results of experiments by Arend[13] have demonstrated that at least at threshold, there is a greater decline in sensitivity to high spatial frequencies than to low spatial frequencies as the stimulus duration is decreased from continuous to 20 msec. Thus the distortion in the spectrum of the input symbol image is due to the temporal response variation among the spatial frequency channels as well as the interference of the uniform field. Similar spatio-temporal effects are being investigated by our laboratory, and should supply a firm base on which to build a model of human performance that can be applied to dynamic as well as static imagery.

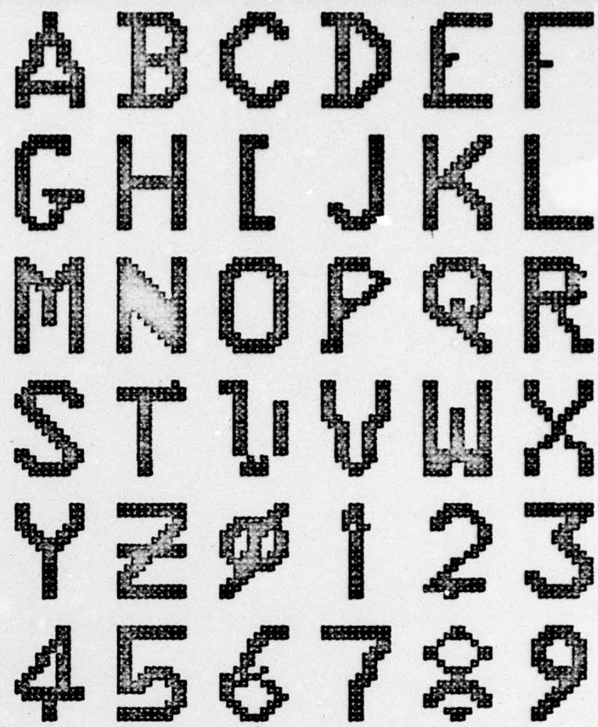
APPENDIX



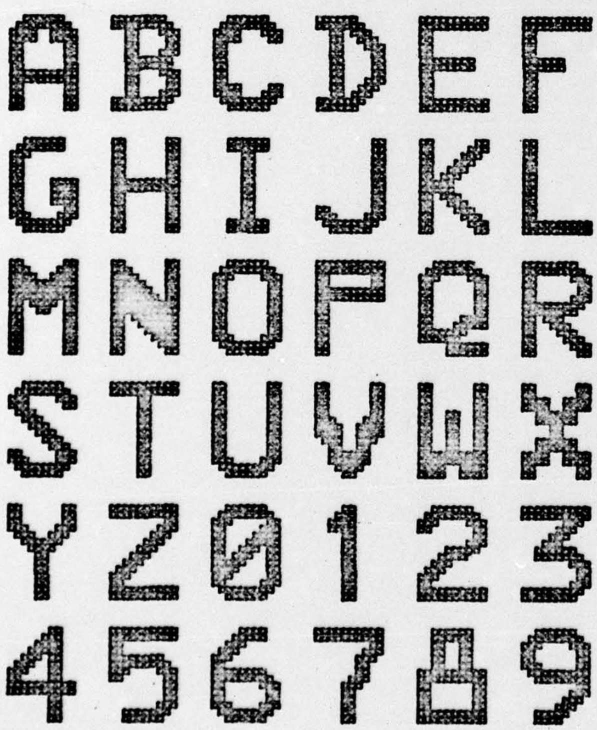
Set 1



Set 2



Set 4



Set 3

References

[1] F. .W. Campbell and J. G. Robson, "Application of Fourier analysis to the visibility of gratings," *J. Physiology (London)*, vol. 197, pp. 551-566, 1968.

[2] F. W. Campbell and D. G. Green, "Optical and retinal factors affecting visual resolution," *J. Physiology (London)*, vol. 181, pp 576-593, 1965.

[3] C. Blakemore and F. W. Campbell, "The existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images," *J. Physiology (London)*, vol. 203, pp. 237-260, 1969.

[4] J. R. Hamerly, R. F. Quick and T. A. Reichert, "A study of grating contrast judgement," *Vision Research*, vol. 17, pp. 201-207, 1977.

[5] C. R. Carlson, R. W. Cohen and I. Gorog, "Visual processing of simple two-dimensional sine-wave luminance gratings," *Vision Research*, vol. 17, pp. 351-358, 1977.

[6] M. Kabrisky, *A Proposed Model for Visual Information Processing in the Human Brain*, Urbana IL and London: University of Illinois Press, 1966.

[7] M. Kabrisky, et al., "A theory of pattern perception based on human physiology," *Ergonomics*, vol 13, pp. 129-142, 1970.

[8] C. Radoy, "Pattern recognition by Fourier series transformations," MSEE Thesis, GE/EE/67-a-11, Air Force Institute of Technology, Wright-Patterson AFB, Ohio, 1967.

[9] O. H. Tallman, "The classification of visual images by spatial filtering," Ph.D. Dissertation, Air Force Institute of Technology, Wright-Patterson AFB, Ohio, CFSTI AD 858-866, 1969.

[10] G. Sperling, "A model for visual memory tasks," in *Information Processing Approaches to Visual Perception*, R. N. Haber, Ed., New York, Chicago: Holt, Rinehart and Winston, Ch. 1, pp. 18-31, 1969.

[11] J. F. Mackworth, "The relation between the visual image and post-perceptual immediate memory," *Journal of Verbal Learning and Verbal Behavior*, vol. 2, pp. 75-85, 1963.

[12] M. I. Posner, "Short-term memory systems in human information processing," *Acta Physiologica*, vol. 27, pp. 267-284, 1967.

[13] L. E. Arend, Jr., "Response of the human eye to spatially sinusoidal gratings at various exposure durations," *Vision Research*, vol. 16, pp. 1311-1315, 1976.